



Docket No.8964-000004/US

IN THE U.S. PATENT AND TRADEMARK OFFICE

Applicant(s): Huai-Jen TSAI

Application No.: 10/632,794

Filed: August 4, 2003

For: NOVEL TRANSGENIC ZEBRAFISH, GENE FRAGMENTS
 AND METHODS FOR PRODUCING TRANSGENIC
 ZEBRAFISH

RULE 32 DECLARATION OF HUAI-JEN TSAI



Declaration

I, Huai-Yen Tsai, have a lot of experience in transgenic techniques since 1998. Based on my experiences of manipulating transgenic techniques, it's not straight forward to predict or assume that how good an engineered construct is expressed in the target organism. Each construct has its unique characteristic of molecular structure. In addition, many biological factors, such as picking a suitable gene fragment, expressional level, mRNA stability, species compatibility, and other biological unknown factors, should be taken into consideration on producing a transgenic fish line that enables to transcribe the foreign gene properly and results in an excellent fluorescent image of whole fish. It cannot be derived from the citation of EG producing green fluorescent fish obviously and directly. The recipient of fish species that are used to carry on transgenesis and the donor of invertebrate species that are supplied the fluorescent protein gene are totally involved in determining whether the excellent glowing transgenic fish can be generated. For example, EGFP gene produced from jelly fish (produced by Clontech Co.) is well expressed in Japanese medaka, but it is not quite prominently expressed in zebrafish. On the other hand, DsRed gene from coral (Clontech) is expressed much more efficiently in zebrafish than in medaka.

Based on my understanding, it's not the case that simply using a red fluorescent protein in substitution for the EGFP in the cited reference can easily produce a red fluorescent fish. Actually TAIKONG CORP. and I had spent many years trying various gene fragment constructions and then finally found that only WAUED is ideal for producing red fluorescent fish. Other alternatives resulted in either high death rate of embryos, or the embryos failed surviving to a mature adult fish. Given the above, please kindly re-judge about the un-obviousness of the invention.

The undersigned hereby declares that information supplied is true in substance and in fact.

Signed at Institute of Molecular and Cellular Biology, College of Life Science,
National Taiwan University, Taipei, TAIWAN

The 25th Day of 14, 2006

Signature

Named Printed in Block Letters
HUAI-JEN TSAI



Copy number related transgene expression and mosaic somatic expression in hemizygous and homozygous transgenic tilapia (*Oreochromis niloticus*)

M. Azizur Rahman, Gyu-Lin Hwang, Shaharudin Abdul Razak, Frédéric Sohm & Norman Maclean*

Division of Cell Sciences, School of Biological Sciences, University of Southampton, Bassett Crescent East, Southampton SO16 7PX, United Kingdom

Received 6 January 2000; revised 27 April 2000; accepted 15 May 2000

Key words: β -galactosidase, mosaic expression, transgenic fish

Abstract

Three lines of transgenic tilapia (*Oreochromis niloticus*) fish were generated with a construct containing a *lacZ* reporter gene spliced to a 4.7 kb 5' regulatory region of a carp beta actin gene. All these three lines contain different copy numbers of transgenes and the levels of *lacZ* expression were found to be related to transgene copy number. Mosaic patterns of somatic *lacZ* expression were observed in these three lines which differed between lines but were consistent within a line. We also observed that expression of the reporter gene in homozygous transgenic fish was approximately two-fold greater than in the hemizygous transgenics. Analysis of expression of the reporter gene on a tissue-to-tissue basis demonstrated that *lacZ* expression of the reporter gene in stably transformed fish occurred with variable intensity in different organs and tissues and was also sometimes variable in different cells of the same tissue in G1 and G2 generations of the transgenic lines.

Introduction

Transgene expression level in stably transgenic organisms is affected by many factors, in particular by the promoter (Dorer, 1997) driving the transgene, the copy number of the transgenes (Whitelaw et al., 1992; Knotts et al., 1995) in the genome, and by interaction between transgene and flanking sequence DNA (Fraser and Grosveld, 1998). Moreover, the copy number of the transgene within a line is not always well correlated with the level of transgene expression in that line, due to the complicating effects of integrated multiple copies in concatameric arrays (Henikoff, 1998). Apparent reduction of expected expression levels by concatamers has been attributed to DNA methylation (Collas, 1998), or formation of heterochromatin (Manuelidis, 1991), or both (Dobie et al., 1997).

In this study we have used a construct containing a *lacZ* reporter gene driven by a carp β -actin promoter (*C β AlacZ*), the latter regulatory sequence extending to 4.7 kb (Rahman et al., 1997). Since the cyto-skeletal β -actin is a house-keeping gene and is expressed ubiquitously in all cells and tissues (Moav, 1994), it is of interest to monitor levels of *lacZ* expression both temporally and in different cells. Despite the ubiquitous expression characteristics of the β -actin gene promoter, ubiquitous or widespread expression was not observed when a range of β -actin/*lacZ* fusion gene constructs were used in lines of transgenic mice, for example human β -actin/*lacZ* (Nillson & Lendahl, 1993) and chicken β -actin/*lacZ* (Sands et al., 1993). Moreover, restricted expression of *lacZ* was observed even when the *lacZ* gene was integrated into the endogenous β -actin locus by homologous recombination (Shawlot et al., 1998). This inappropriate expression of *lacZ* has been found to be widespread when *lacZ* fusion constructs using a regulatory sequences from a ubiquitous gene were tested. This phenomenon

* Author for correspondence: Tel.: +44 2380 594403; Fax: +44 2380 594269; E-mail: nm4@soton.ac.uk

seems to be more common when using a ubiquitous gene promoter than a tissue specific promoter (Cohen-Tannoudji et al., 2000).

Several attempts have been made to produce 'blue mice' using a *lacZ* reporter gene (Cohen-Tannoudji et al., 2000). Only one out of seven transgenic lines containing a HMGCR (3-hydroxy-3-methyl-glutaryl-coenzyme A reductase) gene promoter showed ubiquitous expression (Tam & Tan, 1992). Several other attempts to produce 'blue mice' using a *lacZ* fusion construct containing ubiquitously expressed genes, including HPRT, β -actin, β -macroglobulin and class 1H-2K (cited in Cohen-Tannoudji et al., 2000), have been unsuccessful.

In other studies, when widely expressed MHC class 1H-2K gene regulatory sequences were fused to other reporter genes such as hGH, *c-myc* and *v-jun*, ubiquitous expression in adult organs of transgenic mice have been obtained. However, when the same regulatory sequences fused to *lacZ* were used, the widespread expression of *lacZ* was not observed (cited in Cohen-Tannoudji et al., 2000). These experiments suggest that *lacZ* sequences exert a dominant negative effect over the regulatory sequence of the ubiquitously expressed genes. It has also been observed that incomplete transcription of the *lacZ* gene occurs in yeast cells when the reporter is driven by a GAL1 promoter whereas the full length of mRNA of the PH05 yeast gene was detected using the same promoter in yeast cells (Chávez & Aguilar, 1997).

The present study examines the transgene dosage effect on levels of transgene expression. Our chief finding is that expression levels correlate quite closely with transgene copy number both when fish of the three lines are compared, and when expression levels in homozygous transgenics are compared with those in hemizygous transgenics. However, as has been found in mice in the work mentioned above, there is also evidence for heterocellular or mosaic expression of the *lacZ* gene (Robertson et al., 1996) when the levels of transgene expression are compared from one cell to another in certain tissues of our transgenic fish.

Materials and methods

Generation of fish lines

Three lines of transgenic tilapia were previously produced by micro-injecting 10^5 copies of a construct (*C β AlacZ*) containing a carp β -actin promoter spliced

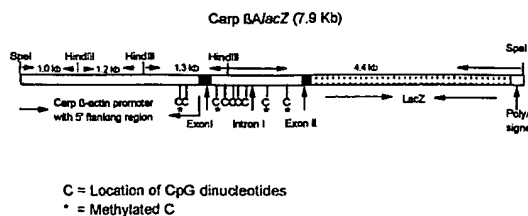


Figure 1. Diagram showing the transgene construct used to produce the three lines of transgenic fish. The 7.9 kb fragment shown was isolated from the plasmid by digestion with *SpeI* and used for injection in this linear form. Restriction sites for *HindIII* are shown and known CpG dinucleotides (C) and CpG sites believed to be methylated in some of our fish (C*) are indicated.

to a *lacZ* reporter gene (Figure 1) into the fertilized eggs of tilapia fish (*Oreochromis niloticus*), as described in Rahman et al. (1997). It should be noted that, as indicated in the paper cited, these fish were produced by coinjection of this construct and another different growth hormone expressing construct. Although both were integrated in the lines of fish discussed here, the second construct does not to our knowledge interfere with the integration patterns reported here. Successive generations of tilapia were produced by crossing a transgenic tilapia with a wild type tilapia. Homozygous fish were produced by crossing two hemizygous transgenic fish in the same line and checking by Southern blotting and subsequent expression in transgenic progeny fish. All the embryos used in this experiment were grown and analysed in contained laboratory conditions.

Southern blotting

Southern blotting was carried out with 3–10 μ g of genomic DNA from fin clips of fish from transgenic tilapia lines. The DNA was digested with the restriction enzyme *HindIII* and blotted onto a nylon membrane. Hybridization was carried out using whole *C β AlacZ* as a probe. For the methylation study, the DNA was digested with either *MspI* or *HpaII* before being blotted onto the membrane. The hybridization in the methylation study was done using the carp beta actin regulatory sequence as a probe. The procedure for the blotting and hybridization of membranes was followed as described in Rahman et al. (1997).

Densitometric determination of gene copy number using Southern blot

Copy numbers were determined by densitometric determinations on a Southern blot (data not shown). The

bands from 'internal' fragments of 1.2 and 1.3 kb were used for the densitometric determination. In a previous paper (Rahman et al., 1997) the copy number estimation was carried out on a single densitometric scale. Since the difference in signal intensity between transgenic lines are too great, accurate estimation of the transgene copy number was difficult. To establish more accurate estimation of transgene copy number and to ensure that determinations were made in the linear part of the densitometric reading, band intensities of standard lanes were compared with transgenic fish DNA using appropriate settings to ensure that readings were always taken in an optimal region of the scale.

Detection of transgene expression

Protein and methyl umbelliferone galactoside (MUG) assay

Embryos aged 4–7 days after fertilisation were homogenized in 60 μ l of homogenisation buffer (10 mM Tris-HCl, pH 7.5; 10 mM NaCl and 0.1% TritonX-100) and then frozen immediately in liquid nitrogen and stored at -20°C until analysis. The samples were incubated at 37°C after thawing on ice and then centrifuged at 14,000 rpm (Eppendorf 241C) for 10 min at 4°C . For the protein assay, 5 μ l of homogenate from an individual sample was taken up into a separate eppendorf tube and the rest of the sample used for MUG assay. The concentration of protein in the supernatant from individual embryo homogenates was determined by the method of Bradford (1976) using bovine serum albumin (BSA) as a standard.

For protein assay analysis, the volume of supernatant was taken to standardise the protein concentration of each embryo for the MUG assay. The MUG assay was carried out by the method described on Hoefer's Technical Bulletin # 129 by using a Hoefer TKO 100 minifluorometer (Braell, 1991). In the MUG assay, 4-methyl umbelliferyl- β -D-galactoside was used as a substrate which was then cleaved by β -galactosidase to yield the fluorescent molecule 4-methylumbelliferone (7-hydroxy-4-methylcoumarin, MU). After excitation at 365 nm, 4-methylumbelliferone emits light at a wavelength of 460 nm. This assay is sufficiently sensitive to detect picogram quantities of β -galactosidase.

X-Gal staining

The embryos and the tail part of the fry were fixed in freshly made fixative (2% paraformaldehyde, 0.2% of 25% glutaraldehyde, 0.02% NP40 and dissolved in

0.1 M phosphate buffer at pH 7.6) for 5–6 h and then incubated in X-gal solution overnight at 30°C . The X-gal stained samples were then washed three times for 5 min each in phosphate buffered saline and photographed. The fixative and X-gal staining procedures were followed as described in Rahman et al. (1997). The same protocol was followed to stain the cell cultures. After X-gal staining, the cells were stained with haematoxylin 0.2% and eosin 0.5% and washed before observation under an inverted phase microscope.

Sectioning of embryos

The X-gal stained embryos were dehydrated in ascending concentrations of ethanol for 5 min at each concentration (30%, 50%, 70%, 80%, 90% and finally twice in 100%). The ethanol in dehydrated embryos was then replaced with three changes of butanol-1-ol and left overnight, and then replaced with three changes of paraffin wax, paraplast plus (Sigma) and left overnight at 58°C in a molten state. The following day the embryos were embedded in fresh paraffin wax and subsequently sectioned at 10 μ m thickness using a rotary microtome. The sections were placed on silicone coated slides (Sigma) and air dried overnight. The dried sections were immersed twice for 10 min each in histoclear and then rehydrated in descending concentrations of ethanol (5 min each time) and finally in distilled water before counter staining in nuclear fast red for 5 min. The slides containing the sections were then washed in distilled water and passed through ascending concentration of ethanol (5 min each) and finally in histoclear for 5 min. The slides were mounted in DPX, a cover slip was added and the slides were left overnight to air dry.

Preparation of the cells

The cell cultures were developed from the tail fins of tilapia, following a protocol modified from Alvarez et al. (1991). Before dissection, the fishes were anaesthetised. 1–2 cm² of tail fin was cut and washed in 1% SDS diluted in a saline buffer (NaCl 100 mM, Na₂HPO₄ 80 mM, NaH₂PO₄ 25 mM, pH 7.4) for 20–30 min. The tail fin was then rinsed twice in the saline buffer, once in 70% ethanol and once in PBS. The tail fin was cut into 2 \times 2 mm pieces. The tail fin pieces were placed in a 100-mm petri dish with complete medium {(Leibovitz L15 (Sigma), foetal calf serum 20% (ICN), antimycotic antibiotic solution 1X (Sigma, A4668) RPMI 1640 amino acid solution 1X (Sigma, R7131), glutamine 4 mM (Sigma,

G7513)). The culture plates were sealed with parafilm and kept at 25°C without CO₂. Half of the medium was changed every two or three days until the cells began to grow from the tail fin pieces. The development of these primary cultures takes between 21 and 28 days. The cells can be stained at this point or transferred onto a new plate for further growth.

Results

Concatamerization and integration of transgenes

From the densitometric measurements of the Southern blot (data not shown), it is clear that multiple copies of the transgene were integrated into these three lines (C58, C86 and C118 lines). In most cases it is observed that most transgenes are ligated in a head-to-tail arrangement as compared to those showing an inversely repeated sequence. The internal fragment bands of 1.2 and 1.3 kb were used for this calibration. It is estimated that these three lines contain approximately 19, 34, and 330 copies of *CβAlacZ* transgene in the C86, C118 and C58 lines, respectively. These figures for copy number are a revision of figures previously published by us in Rahman et al. (1997). Use of a revised method has permitted more accurate densitometric analysis.

Tissue specific expression and non-uniform expression of transgenes in hemizygous and homozygous fish

Since in these three transgenic lines the reporter *lacZ* gene is driven by a carp β -actin gene regulatory sequence, it might be expected that expression of *lacZ* would be observed in all cells and tissues of the transgenic fish. However, the *in-situ* X-gal staining of the embryos reveals that the expression of *lacZ* is somewhat patchy in G1 and G2 individuals of these three lines (Figure 2 A–E). Thus, although the intensity of X-gal staining was found to roughly correspond to copy number in all lines, there was some variation in the tissue specific levels of expression. This is presumably a result of variable activity of the beta actin promoter in different tissues. Thus parts of the brain such as olfactory lobe invariably show staining, whereas the mid brain shows little or no staining (see Figure 2C, D and E). Analysis of whole-mount, stained embryos revealed that the beta-galactosidase expression patterns were transmitted with little variation to successive generations in a particular line. Paraffin embedded sections were made from G1/G2

embryos from these three lines and the expression of *lacZ* was examined either on a cell by cell or tissue by tissue basis. From the analysis of sections it was observed that *lacZ* expression was found to be present in most of the tissues (except in the C86 line) in a mosaic fashion, that is, that some cells in a tissue failed to express *lacZ* or expressed at a low level at a given time, as summarised in Table 1. This is particularly clearly indicated in muscle sections of the C86 line (see Figure 2F).

Homozygous progeny were produced by crossing two G2 hemizygous transgenic individuals from the C118 line, when 25% of the progeny is expected to be homozygous for the transgene. From X-gal staining it was observed that approximately 25% of the progeny showed a much higher intensity of X-gal staining compared to the hemizygous transgenics. The whole mount X-gal staining of the homozygous progeny revealed that the pattern of expression in hemizygous fish is similar to the homozygous but the amount of *lacZ* expression is much higher on a tissue-to-tissue basis in homozygous as compared to hemizygous (Figure 2 H and Figure 3).

In summary, we find evidence for two particular factors determining differential expression. The first is indicative of variable promoter activity in different tissues and times (see, for example expression in different portions of brain in Figure 2C, D and E).

Figure 2. (A) Section paraffin embedded 5 days post-fertilization G1 embryo of the C118 line (magnification 32 \times). (B) Section of paraffin embedded 9 days post fertilization G1 embryo of the C58 line (magnification 25 \times). Both embryos were fixed, then subjected to X-gal staining, then embedded, and finally counter stained with nuclear fast red, post embedding. (C), (D) and (E) are brains of 7 months old G2 fish of the C86, C118 and C58 lines, following dissection, and stained after removal, with X-gal (magnification 10 \times). Note the strong expression in one area of the telencephalon (left of the picture) in all three lines, contrasting with the limited expression in the cerebellum and optic tectum evident only in the C58 line (E). (F) and (G) are whole mount specimens of muscle dissected from hemizygous fish of C86 and C118 lines respectively stained with X-gal. Note, particularly in 5F, the fact that some muscle fibres show strong expression while many adjacent fibres show complete absence of expression (magnification 10 \times) (Reproduced with permission from Razak SA, Hwang G-L, Rahman MA, Maclean N (1999). Growth performance and gonadal development of growth enhanced transgenic tilapia *Oreochromis niloticus* (L.) following heat-shock-induced triploidy. *Mar. Biotechnol* 1: 533–544). H is a photo of X-gal stained whole mount embryos of 9 days post fertilization resulting from a cross between sperm and eggs of hemizygous C118 line. A ratio of 25% non expressing, 50% hemizygous expressing, and 25% homozygous expressing, is expected. Examples of all three classes are visible. (magnification 10 \times).

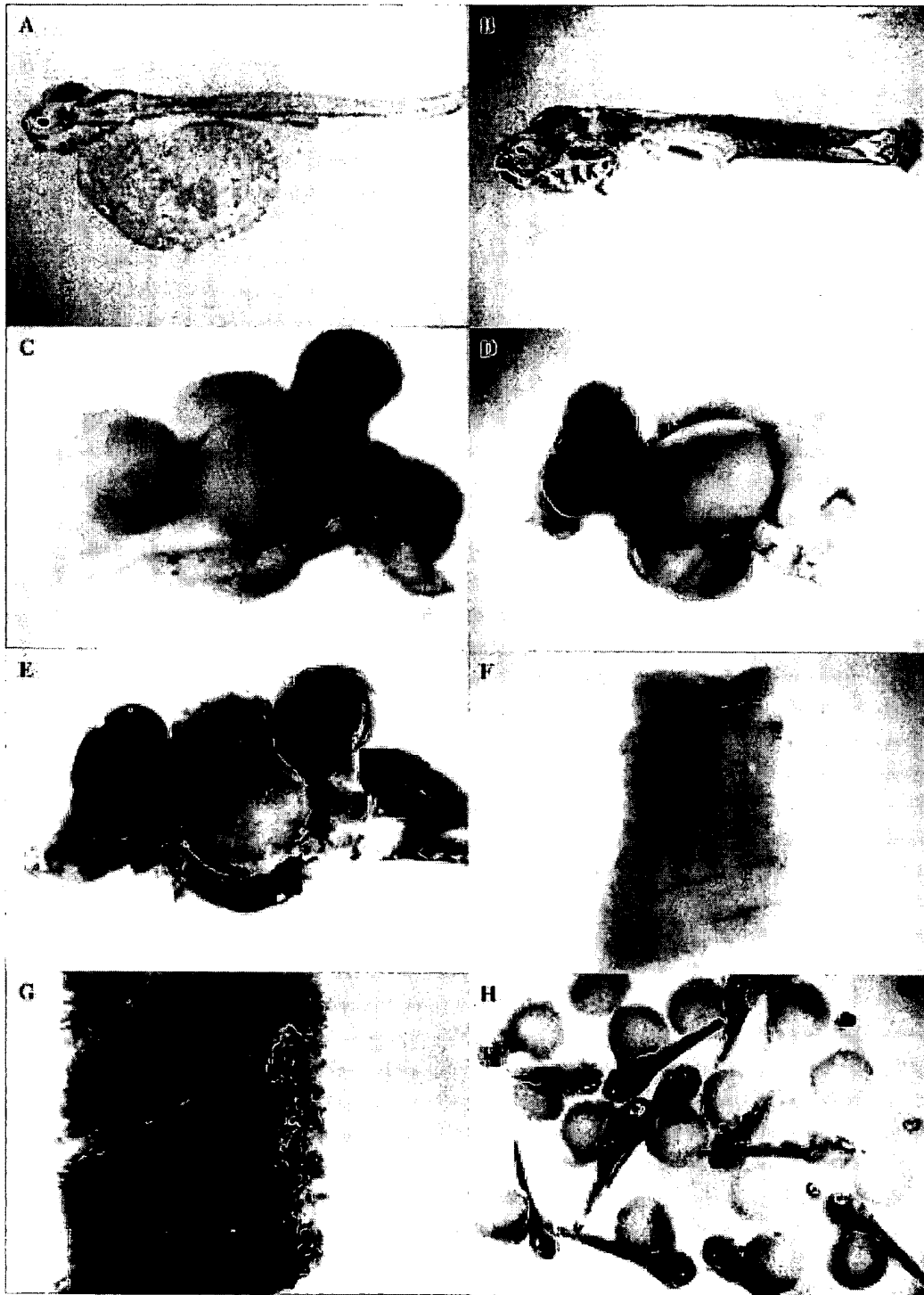


Figure 2.

Table 1. Copy number and expression characteristics of transgenic fish in three transgenic lines

| Transgenic lines | Copy number | Amount of lacZ expression (10 ³ × MU units/μg protein) Average ± Standard deviation | | | | lacZ expression observed in different tissue section of transgenic G2 embryos. n = 18 | | | | | | | Level of LacZ expression (10 ³ × MU units/μg protein. Average ± Standard deviation | | | |
|------------------|-------------|--|---------------------|----------------------|-----------------------|--|-----|----|----|-----|-----|-----|---|----------------------|----------------------|----------------------|
| | | 4d | 5d | 6d | 7d | BR | YS | MS | GL | SC | KD | HT | Stained embryos | | 4-week old fry | |
| | | | | | | | | | | | | | HEM | HOM | HEM | HOM |
| C86 | 19 | 0.6 ± 0.1 n = 12 | 1.2 ± 0.2 n = 7 | 1.8 ± 0.3 n = 5 | 5.5 ± 1.6 n = 13 | +/- | +/- | +/ | - | - | - | - | | | | |
| C118 | 34 | 1.4 ± 0.3 n = 10 | 3.5 ± 0.4 n = 7 | 4.2 ± 1.6 n = 10 | 10.9 ± 1.7 n = 12 | +/ | +/- | +/ | +/ | +/- | +/- | - | 41.9 ± 11.0 n = 5 | 81.9 ± 31.2 n = 5 | 58.4 ± 7.9 n = 23 | 91.5 ± 12.4 n = 9 |
| C58 | 330 | 3.6 ± 2.4 n = 12 | 9.6 ± 8.4 n = 9 | 67.7 ± 45.0 n = 8 | 76.6 ± 44.4 n = 13 | +/ | +/ | +/ | +/ | +/ | +/ | +/- | | | | |
| Control | - | 0.3 ± 0.1 n = 10 | 0.7 ± 0.2 n = 10 | 0.8 ± 0.1 n = 10 | 3.1 ± 1.1 n = 10 | - | - | - | - | - | - | - | | | | |

BR = Brain; YS = Yolk Sac; MS = Muscle; GL = Gills; SC = Spinal cord; KD = Kidney; HT = Heart.

HEM = Hemizygous; HOM = Homozygous. n = No. of individuals

- = No expression; + = low expression; ++ = Moderate expression; +++ = High expression.

Presence of both - and + signs indicate that some individual express and some do not express the transgene in that particular tissue.

The second is apparent mosaic expression in at least some lines when adjacent muscle fibres are compared in Figure 2F. However, we have no evidence to confirm that this very patchy expression pattern is a result of chromatin interactions, nor does it occur only in hemizygous transgenics. It may be an anomaly of the activity of the transgene construct used, and the time of beta actin synthesis in the cell cycle. Thus we find in primary tissue cultures set up from the fin clip cells of fish of lines C58 and C86 that some cells do not express, but there are no obvious clones of expressing or non-expressing cells (Figure 4).

Copy number and level of expression in transgenic lines

In order to correlate the level of expression with transgene copy number, the expression of *lacZ* was determined fluorometrically using the MUG assay in G3 progeny of the C86 and C118 lines and G2 progeny of the C58 line for four successive days beginning from four days after fertilisation. As mentioned earlier in the context of the mosaic expression in individuals in these three lines, the level of expression was found to vary slightly among individuals in the same line. It is interesting to compare the expression levels of the 7 day-old embryos of each line with the transgene copy number calculated by densitometry of the Southern blot. Line C86, with a presumed copy number of 19, shows an average expression intensity of 5.51; line C118, with a presumed copy number of 34, shows an average expression level of 10.93, and line C58, with a presumed copy number of 330, shows an average expression level of 76.61 (Table 1, Figure 5). If the expression levels are divided by the presumed copy number, figures for average expression level per gene copy are 0.29, 0.32 and 0.23. These figures indicate a rather good correlation between transgene copy number and expression level. Note that the expression levels of the co-injected growth hormone coding sequence could not be compared, since radioimmunoassay was carried out only on fish of one of the lines.

The levels of expression were also compared between the hemizygous and homozygous transgenic individuals amongst 4-week old fry. In this analysis, homozygous individuals were identified by X-Gal staining of the tail part of the fry, which showed higher staining intensity than did the hemizygous (Figure 3). The remaining part of the body was in each case analysed quantitatively using the MUG assay for

measuring the level of *lacZ* expression. DNA analysis by Southern blotting was also carried out on samples from this material. From the MUG assay it was observed that although the expression level was found to vary from individual to individual in both groups, the expression level in the homozygous fish was on average approximately two times higher than the expression level of hemizygous transgenic fish (Table 1). The Southern blot analysis also confirmed that the homozygous transgenic individuals contained approximately twice the amount of transgenic DNA in their genome (Data not shown). The same membrane was rehybridized with an endogenous DNA probe to confirm that the same amount of DNA was loaded in each lane. The endogenous DNA probe was obtained by digesting control fish DNA with *HindIII* enzyme and recovering an approximately 1 kb size band from the agarose gel. This DNA was then subsequently purified.

Methylation of the transgenes

Fin DNA from three adult tilapia from each of the three transgenic tilapia lines (except in the C58 line where one adult and two juveniles (2-months old) were used due to an insufficient number of adult fish) were analyzed to determine the methylation status of the transgenic DNA. The isoschizomers *MspI* and *HpaII* were used to observe the methylation pattern in the carp β -actin regulatory sequences using only the promoter plus other 5' regulatory sequences as a probe since the *MspI* enzymes cuts regardless of methylation, and *HpaII* does not. Analysis of the Southern blot reveals that most of the CpG sites in the transgene are methylated in all three transgenic lines (Figure 6) except in lanes L and N which contained DNA from juvenile tilapia from the C58 line. The presence of 3.3 kb, 540 bp, 241 bp and 214 bp bands in *MspI* digest and absence of such bands in *HpaII* digest suggests that relevant CpG sites are methylated in most transgene copies in these three lines. The approximately 6 kb band is presumed to represent the flanking sequence of transgene present in all *MspI* digested lanes as also is the one extra band in close proximity to that flanking sequence band present in DNA from three fish of the C58 and C86 lines (lanes I, K, M, O, Q and S), but absent in the C118 line (lanes C, E and G). The *MspI* digested band patterns in three individuals of each tilapia line are identical but different between lines, which suggests that the same CpG sequences in the transgene in different lines are differentially

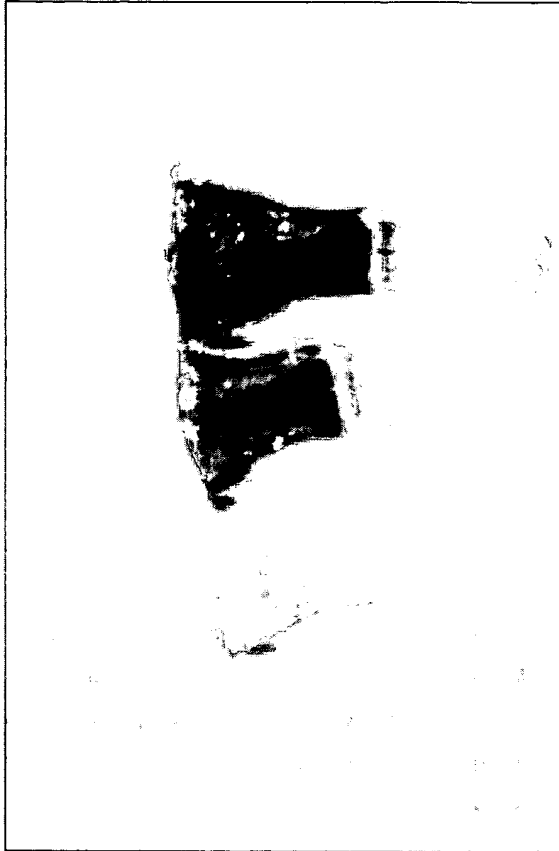


Figure 3.

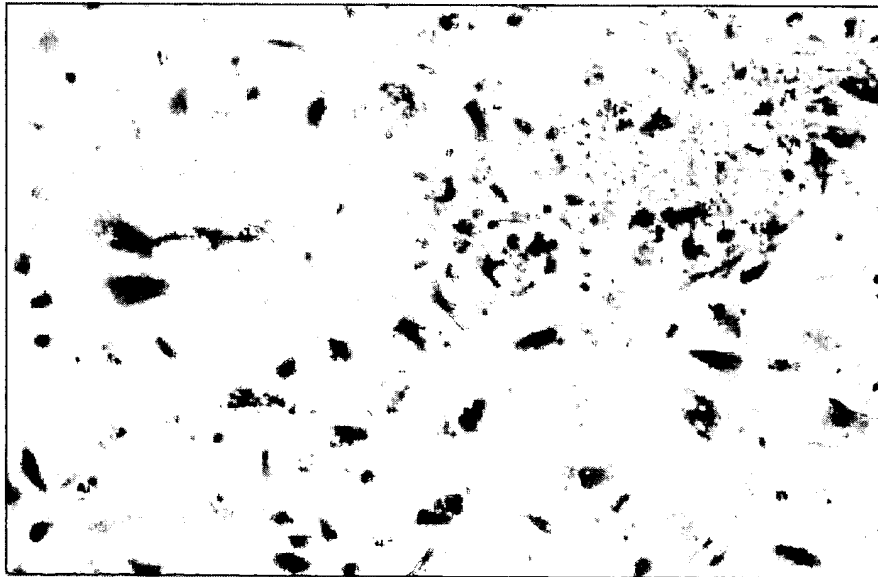


Figure 4. Photomicrography of hemizygous C58 tail fin cell culture. The X-gal staining allows specific identification of cells from C58 fish (which contain the highest number of copies of the transgene) and most of the cells are strongly stained. The level of β -galactosidase expression is roughly constant between different cultures derived from the same fish line but is highly different between cultures derived from different fish lines. (magnification $\times 200$).

Figure 3. Comparison of X-gal staining in tail sections of presumptive hemizygous and homozygous fish from a cross between sperm and eggs of C118 line. A presumptive homozygous transgenic is at the top, hemizygous transgenic in the middle and homozygous wild type at the bottom. (magnification $10\times$).

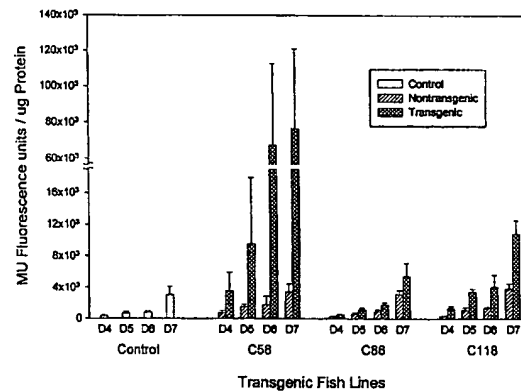


Figure 5. Histograms of levels of *lacZ* expression in embryos of different ages and lines. Expression is determined by the MUG fluorescence assay and the age of the embryos post fertilization (in days) is indicated as D4, D5 etc. The numbers of embryos in each sample are indicated in Table 1, varying between 5 and 13. The large variability between expression levels of C58 embryos on days 6 and 7 as compared to day 5 is conspicuous.

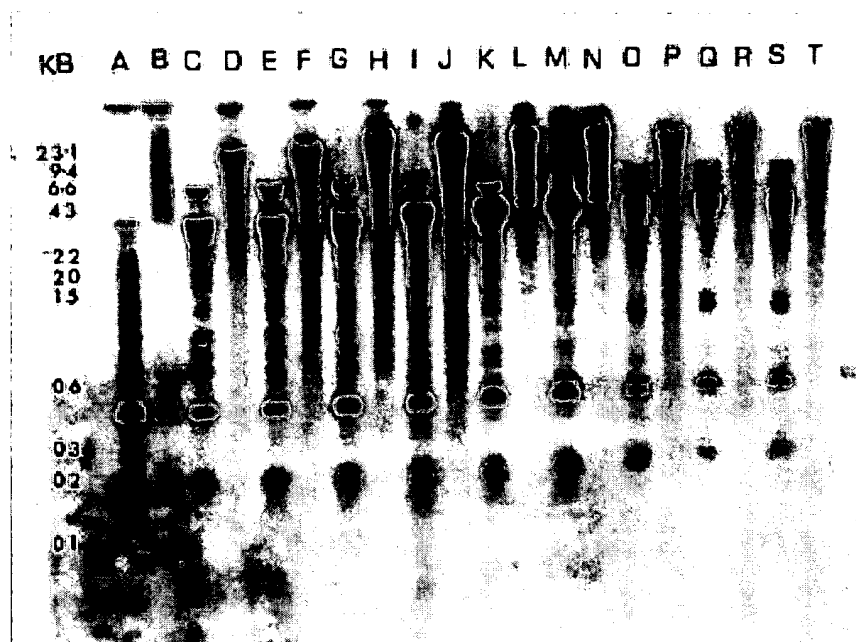


Figure 6. Southern blot analysis of fin DNA from G1/G2 transgenic fish from C58, C86 and C118 lines. Three microgram, 5 μ g and 10 μ g of DNA from fish of C58, C118, C86 line, respectively, were digested with either *MspI* or *HpaII* and hybridized with carp β -actin promoter with 5' sequences. Consecutive pairs of lanes contain the same DNA but the first lane was digested with *MspI* and the following lane digested with *HpaII*. Lanes A–B: 7.9 Kb *C β AlacZ* construct; Lanes C–H: DNA from 3 adult fish from C118 line; Lanes I–N: DNA from 3 adult fish from C58 line; Lanes O–T: DNA from 3 adult fish from C86 line.

methyated. Southern blot analysis thus indicates that the transgene in all three lines is extensively methylated and there is no direct relationship between the increase in copy number and the level of transgene methylation.

Discussion

In this study, three transgenic lines of tilapia were generated using a 7.9 kb transgene with 4.7 kb of 5' flanking sequences. All of these three lines express the transgene in most tissues examined. The *lacZ* expression level in these transgenic lines showed an expression level approximately in line with the transgene copy number, which suggests that the 4.7 kb 5' sequence has been effectively insulated, and the promoter and each transgene in a concatamer can function as an independent regulatory unit (Knotts et al., 1995). Work with other transgenic systems indicates that repeat induced gene silencing (RIGS) is common in plants (Assaad et al., 1993) and has also been observed in transgenic mice (Manuelidis, 1991), where copy number dependent silencing has been observed.

Thus copy number dependent gene expression is the exception rather than the rule in transgenic organisms. The degree of silencing has been found to be increased in inverted repeat transgene arrays, perhaps since inverted repeats are able to form hairpin structure which pair more easily than looped structures (Dobie et al., 1997). In these three transgenic fish lines, most transgene copies were ligated in a head to tail array and we have found little evidence for inverted repeat copies of transgenes. Therefore, it is perhaps unsurprising that the level of *lacZ* expression in these lines was found to be increased in magnitude as transgene copy number increased. In this study when the level of *lacZ* expression in the transgenic embryos was measured on four successive days after hatching, we find that increasing levels of *lacZ* expression were observed with time each day (Table 1 and Figure 5).

The expression levels of homozygous transgenic embryos were analysed by both X-Gal staining and also quantitative expression MUG assay and it was found that approximately twice the amount of *lacZ* expression was detectable in homozygous as compared to hemizygous embryos, which suggests that both

alleles were expressed independently. Co-suppression of transgenes has been observed with homozygous transgenes in transgenic plants (de Carvalho et al., 1992) where the level of a particular RNA may act as a trigger for subsequent RNA degradation (Smith et al., 1994).

In our three transgenic lines, the expression of the *lacZ* reporter gene is regulated by a carp β -actin gene regulatory sequences. As a 'house-keeping gene' this regulatory sequence would be expected to drive *lacZ* expression in all tissues. In this study it was observed that some individuals were indeed able to express in most tissues examined, as for example in C118 and C58 lines, whereas in the C86 line expression of *lacZ* could not be observed in most tissues (Table 1). However, this is most probably a result of the sensitivity threshold of the assay. In mice, five out of five transgenic lines harbouring a 'house-keeping gene' were analyzed and showed extreme variation in tissue levels of expression in the same transgenic line (Mehtali et al., 1990). The presence of CpG islands is a common feature of house-keeping genes, and some of these islands occur in the regulatory regions of the genes (Cohen-Tannoudji et al., 2000). The very high content of C+G and CpG dinucleotides in the *lacZ* gene may act as a substrate for endogenous methylases (Thorey et al., 1993). Therefore, *lacZ* may interfere with the protective mechanism in the CpG islands by methylation of CpG islands in the regulatory region which might cause the silencing or erratic expression of *lacZ* (Cohen-Tannoudji et al., 2000). The aberrant expression of *lacZ* may therefore be due to low numbers of Sp1 transcription binding sites in *lacZ* gene since SP1 sites have been shown to protect against *denovo* methylation (Cohen-Tannoudji et al., 2000). It could also result from incomplete transcription of the prokaryotic gene in a eukaryotic system where, in some cases, most of the *lacZ* mRNA has been found to be degraded within 30 min (Chev  s & Aguilera, 1997).

The patterns of patchy or mosaic expression of *lacZ* in different generations are found to be consistent within the lines and no obvious cessation of *lacZ* expression was observed in G3 and G4 generations of transgenic lines (data not shown). This patchy expression could be due to silencing of *lacZ* gene expression in some clonal cells derived from the cells where the transgene may be methylated (Dobie et al., 1997). In this study, there is no direct correlation between the transgene copy number and the methylation status observed in these three transgenic lines. All fin DNAs

from adults of three lines were found to be methylated except DNA from two juvenile tilapia (Figure 6 Lanes L and N) of the C58 line (which contains the highest transgene copy) where the CpG site in the promoter of some transgene copies is not methylated. Although expression of *lacZ* was detected in the fins of early embryos, lack of *lacZ* expression was observed in adult fins of transgenic fish of these three lines (data not shown). Therefore, methylation of the transgene perhaps not only depends on transgene copy number, but could also be an age related gene silencing mechanism as observed in transgenic mice where the extinction of expression of globin transgene increases with the age of the transgenic mice (Robertson et al., 1996). Similar results have also been observed in transgenic mice where a human keratin 18/*lacZ* fusion gene has been shown to provide appropriate expression in embryonic stages but to become aberrant in adult stages (Thorey et al., 1995). This suggests that *lacZ* may not be a reliably neutral reporter gene in all developmental situations.

Acknowledgements

This work was supported by the Department for International Development (DFID) Fish genetics Research Programme as part of attempts to develop cost efficient small-scale aquaculture systems. We thank Arati Iyengar for useful discussion, and three anonymous referees for useful comments and criticism.

References

- Alvarez MC, Otis J, Amores A and Guise K (1991) Short-term cell culture technique for obtaining chromosomes in marine and fresh water fish. *J Fish Biol* **39**: 817–824.
- Assaad FF, Tucker KL and Signe ER (1993) Epigenetic repeat-induced gene silencing (RIGS) in *Arabidopsis*. *Plant Mol Biol* **22**: 1067–1085.
- Bradford MM (1976) A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the protein-dye binding. *Anal Biochem* **72**: 248–254.
- Braell WA (1991) β -galactosidase assay using the TKO 100 mini-fluorometer. Hoefer Scientific Instruments, *Technical Bulletin no* **129**.
- Ch  vez S and Aguilera A (1997) The yeast HPR1 gene has a functional role in transcriptional elongation that uncovers a novel source of genome instability. *Genes Dev* **11**: 3459–3470.
- Cohen-Tannoudji M, Vandormael-Pournin S, Drezen J-M, Mercier P, Babinet C and Morello D (2000) *LacZ* sequences prevent regulated expression of house keeping genes. *Mechanisms Dev* **90**: 29–39.

- Collas P (1998) Modulation of plasmid DNA methylation and expression in zebrafish embryos. *Nucleic Acids Res* 26: 4454–4461.
- de Carvalho F, Gheysen G, Kushnir S, Montagu MV, Inzé D and Castresana C (1992) Suppression of β -1,3-glucanase transgene expression in homozygous plants. *EMBO J* 11: 2995–2602.
- Dobie K, Mehtali M, McCleanagh M and Lathe R (1997) Variegated gene expression in mice. *Trends in Genet* 13: 127–130.
- Dorer DR (1997) Do transgene arrays form heterochromatin in vertebrates? *Transgenic Res* 6: 3–10.
- Fraser P and Grosveld F (1998) Locus control region chromatin activation and transcription. *Curr Opin Cell Biol* 10: 361–365.
- Henikoff S (1998) Conspiracy of silence among repeated transgenes. *BioEssays* 20: 532–535.
- Knotts S, Rindt H and Robbins J (1995) Position independent expression and developmental regulation is directed by the β myosin heavy chain gene's 5' upstream region in transgenic mice. *Nucleic Acid Res* 23: 3301–3309.
- Manuelidis L (1991) Heterochromatic features of an 11-megabase transgene in brain cells. *Proc Natl Acad Sci USA* 88: 1049–1053.
- Mehtali M, LeMour M and Lathe R (1990) The methylation-free status of a housekeeping transgene is lost at high copy number. *Gene* 91: 179–184.
- Moav B (1994) Developmental and molecular genetics in Fish. *Israel J Zool* 40: 441–466.
- Nillson E and Lendahl U (1993) Transient expression of a human beta-actin promoter *lacZ* gene introduced into mouse embryos correlates with a low degree of methylation. *Mol Reprod Dev* 34(2): 149–157.
- Rahman MA, Iyengar A and Maclean N (1997) Coinjection strategy improves integration efficiency of a growth hormone gene construct resulting in lines of transgenic tilapia (*Oreochromis niloticus*) expressing an exogenous growth hormone gene. *Transgenic Res* 6: 369–378.
- Razak SA, Hwang G-L, Rahman MA, Rahman MA, Maclean N (1999) Growth performance and gonadal development of growth enhanced transgenic tilapia *Oreochromis niloticus* (L.) following heat-shock-induced triploidy. *Mar Biotechnol* 1: 533–544.
- Robertson G, Garrick D, Wilson M, Martin DIK and Whitelaw E (1996) Age-dependent silencing of globin transgenes in the mouse. *Nucleic Acid Res* 24: 1465–1471.
- Sands AT, Hansen TN, Demayo FJ, Stanley LA, Xin L and Schwartz RJ (1993) Cytoplasmic beta-actin promoter produces germ-cell and preimplantation embryonic transgene expression. *Mol Reprod Dev* 34: 117–126.
- Shawlot W, Deng JM, Fohn LE and Behringer RR (1998) Restricted beta-galactosidase expression of a hygromycin-lacZ gene targeted to the beta-actin locus and embryonic lethality of beta-actin mutant mice. *Transgenic Res* 7: 95–103.
- Smith HA, Swaney SL, Parks TD, Wernsman EA and Dougherty WG (1994) Transgenic plant-virus resistance mediated by untranslatable sense RNAs—Expression, regulation and fate of non-essential RNAs. *Plant Cell* 6: 1441–1453.
- Tam PP and Tan SS (1992) The somitogenetic potential of cells in the primitive streak and the tail bud of the organogenesis-stage mouse embryo. *Development* 115: 703–715.
- Thorey IS, Meneses JJ, Neznanov N, Kulesh DA, Pedersen RA and Oshima RG (1993) Embryonic expression of human keratin 18 and K18- β -Galactosidase fusion genes in transgenic mice. *Dev Biol* 160: 519–534.
- Whitelaw BA, Harris S, McCleanagh M, Simons PJ and Clark JA (1992) Position-independent expression of the ovine β -lactoglobulin gene in transgenic mice. *Biochem J* 286: 31–39.

Multiple Effects of Genetic Background on Variegated Transgene Expression in Mice

Margaret L. Opsahl,* Margaret McClenaghan,* Anthea Springbett,* Sarah Reid,* Richard Lathe,[†] Alan Colman[‡] and C. Bruce A. Whitelaw*,¹

*Roslin Institute (Edinburgh), Division of Molecular Biology, Roslin, Midlothian, EH25 9PS, United Kingdom, [†]Centre for Genome Research, University of Edinburgh, Edinburgh, EH9 3JQ, United Kingdom and [‡]PPL Therapeutics, Roslin Biotechnology Centre, Roslin, Midlothian, EH25 9PP, United Kingdom

Manuscript received August 28, 2001

Accepted for publication December 14, 2001

ABSTRACT

BLG/7 transgenic mice express an ovine β -lactoglobulin transgene during lactation. Unusually, transgene expression levels in milk differ between siblings. This variable expression is due to variegated transgene expression in the mammary gland and is reminiscent of position-effect variegation. The BLG/7 line was created and maintained on a mixed CBA \times C57BL/6 background. We have investigated the effect on transgene expression of backcrossing for 13 generations into these backgrounds. Variable transgene expression was observed in all populations examined, confirming that it is an inherent property of the transgene array at its site of integration. There were also strain-specific effects on transgene expression that appear to be independent of the inherent variegation. The transgene, compared to endogenous milk protein genes, is specifically susceptible to inbreeding depression. Outcrossing restored transgene expression levels to that of the parental population; thus suppression was not inherited. Finally, no generation-dependent decrease in mean expression levels was observed in the parental population. Thus, although the BLG/7 transgene is expressed in a variegated manner, there was no generation-associated accumulated silencing of transgene expression.

POSITION-effect variegation (PEV) was first characterized in *Drosophila melanogaster* 70 years ago (MULLER 1930). This involved a mutation-induced translocation of the *white* eye color gene from its normal euchromatic site to an integration point close to centromeric heterochromatin. The juxtaposition led to variegated expression of the gene, the physical manifestation being the mottled eye phenotype. In the intervening years PEV has been extensively studied in *Drosophila*, where the combination of numerous characterized strains and short generation time has allowed complex genetical studies. This work has shown that many modifiers of PEV exist (HENIKOFF 1990; SINGH 1994), with the relative abundance of a given modifier or group of modifiers influencing the ratio of silenced to expressing cells for a gene displaying PEV.

Variegated expression patterns have recently been observed in transgenic studies in mammals (McGOWAN *et al.* 1989; DOBIE *et al.* 1996; FESTENSTEIN *et al.* 1996; ROBERTSON *et al.* 1996; GIRALDO *et al.* 1999; SUTHERLAND *et al.* 2000); PEV may occur in mice (CATTANACH 1974). Although in some cases this variegated transgene silencing appears to reflect a PEV-like phenomenon (DOBIE *et al.* 1997), in others it seems to be at odds with

the classical characteristics of PEV. For example, in some reports the extent of variegation remains constant between siblings within a line (ROBERTSON *et al.* 1996). To be precise, in these mice the amount of transgene product or number of cells expressing the transgene remains constant between individuals of the same line. In contrast, we identified a line of transgenic mice (BLG/7) where there were clear within-line differences in the extent of variegation (DOBIE *et al.* 1996). This case is different from the previous because the amount of product or number of cells expressing the transgene varies between individuals (siblings) of the same line. In this line of mice, the β -lactoglobulin transgene is integrated close to the centromere of chromosome 15 as a tandem array of ~ 25 copies. The variegated expression of the transgene in this line therefore exhibits some of the characteristics of PEV (DOBIE *et al.* 1997).

Different genetic backgrounds can have widely differing effects on transgene expression (McGOWAN *et al.* 1989; ELLIOT *et al.* 1995; SCHWEIZER *et al.* 1998), reflecting the presence of strain-specific modifiers of expression. The experimental overexpression of a single heterochromatin protein can modify the extent of transgene variegation in mice (FESTENSTEIN *et al.* 1999; McMORROW *et al.* 2000). BLG/7 was originally created on a mixed CBA \times C57BL/6 background, and it can thus be argued that strain-specific modifiers could affect the observed variegation. However, initial analysis of 3rd generation backcrosses provided no evidence for a

¹Corresponding author: Roslin Institute (Edinburgh), Division of Molecular Biology, Roslin, Midlothian, EH25 9PS, United Kingdom.
E-mail: bruce.whitelaw@bbsrc.ac.uk

major strain-specific modifier of variegation affecting the transgene array at this locus (DOBIE *et al.* 1996). There was, nonetheless, a small trend toward a lower mean expression level for both populations, suggesting that minor modifiers may exist. In view of this observation, we continued the backcrossing regime for 13 generations. This allowed us to investigate the existence of minor modifiers. In addition, the activity of the transgene with respect to inbreeding and generation effects could be assessed.

MATERIALS AND METHODS

Mice: Transgenic mouse line BLG/7 carries ~25 copies of a 16-kb *SalI-SalI* ovine genomic fragment encompassing the milk protein gene β -lactoglobulin (SIMONS *et al.* 1987; WHITEHEAD *et al.* 1992; DOBIE *et al.* 1996) as a tandem array inserted close to pericentric heterochromatin of chromosome 15 (DOBIE *et al.* 1996). The BLG/7 line was generated by microinjection of the transgene into F₁ CBA \times C57BL/6 eggs with subsequent line maintenance through mating transgenic males with F₁ CBA \times C57BL/6 females (termed the parental population). This mating regime has been maintained in parallel with backcross matings generating isogenic lines (termed the backcross populations). Generation of isogenic lines was by mating male transgenic mice with either CBA or C57BL/6 female mice. Outcross mice were generated by mating isogenic C57BL/6 BLG/7 males to female CBA mice, using resulting F₁ females for our analysis. Due to breeding difficulty, the CBA backcross population was not maintained beyond the 13th generation.

Milk collection and processing: Milks were collected at day 11 of lactation and processed as described (McCLENAGHAN *et al.* 1995). Individual 1/250 diluted samples were analyzed on protein gels, 50 μ l of the diluted samples were run for the backcross milks, and 40 μ l of the diluted samples were run for the parental milks.

Total milk protein content was determined by micro-Kjeldahl analysis of pooled milk samples from backcross and parental populations.

Protein gels: SDS/PAGE analysis of milks was performed using 17.5% discontinuous gels (37.5:1 acrylamide:bisacrylamide, Anachem), 1-mm-thick, 20-cm plates with a standard 4% stacking gel (Bio-Rad Protean II rig used), and stained overnight with Coomassie Blue G-250. Purified sheep β -lactoglobulin standards, whose protein content had been determined using the micro-Kjeldahl technique (ASCHAFFENBURG and DREWRY 1957), were run on each gel. Quantification of β -lactoglobulin in transgenic samples (average of three repeats) was by pixel intensity (Multi Analyst Program version 1.1, Bio-Rad, Hercules, CA). Values were normalized for loading using the β -casein of the nontransgenic control.

Statistical analysis: Satterthwaite's approximation was used to calculate the degrees of freedom (d.f.) for a *t*-test between sample means in cases where the sample variances differed significantly. This gives a more conservative test than the standard *t*-test. Coefficients of variation ($100 \times \text{SD}/\text{mean}$) were also used to compare variation between samples. These give a measure of variability as a percentage of the mean value, which is appropriate for samples from populations whose variation tends to increase with the mean.

FISH: Fluorescent *in situ* hybridizations (FISH) were performed on 5- μ m sections from paraformaldehyde-fixed, wax-embedded mammary tissue.

Prehybridization: Sections were deparaffinized using two 20-min and one 5-min washes in xylene and then rehydrated following standard procedures. This was followed by 2 min $1 \times$ PBS, 10 min

4% paraformaldehyde (PFA) in PBS (pH 7.2), two 2 min $1 \times$ PBS, 7.5 min Proteinase K digestion (50 mM Tris, 5 mM EDTA, 5 μ g/ml Proteinase K), 1 min $1 \times$ PBS, 2 min 4% PFA in PBS (pH 7.2), 10 sec H₂O, 30 sec 0.1 M TEA (or TCA), two 5 min 0.1 M TEA (or TCA) with 312.5 μ l/100 ml acetic anhydride, 2 min each $1 \times$ PBS and 0.85% NaCl, and then sections were dehydrated using standard procedures and air dried.

Hybridization: Sections overlaid with buffer containing probes were heated for 6 min at 90°, iced for 1 min, and hybridized O/N at 37°. Sections were washed 10 + 30 min at 37° in $2 \times$ SSC and then 10 + 30 min RT in $0.1 \times$ SSC and mounted using Vectashield mounting medium with antifading agent. Probes were antisense 30' oligos based on cDNA sequences labeled with either Cy3 (BLG) or fluorescein (β -casein). BLG probes were for positions 1621–1650, 1681–1710, 2591–2620, 3851–3880, and 4581–4610 using GenBank sequence X12817. β -Casein probes were for positions 9291–9320, 9411–9440, 9471–9500, 9531–9560, and 9621–9650 using GenBank sequence X13484. Synthesis and labeling were performed by MWG-BIOTECH AG (Ebersberg, Germany). The oligos were combined to a final concentration of 40 ng/ml in the hybridization buffer (40% formamide, $2 \times$ SSC, $1 \times$ Denhardt's, 10% dextran sulfate, 50 mM phosphate buffer, 50 mM dithiothreitol, 0.250 mg/ml tRNA, and 0.5 mg/ml denatured salmon sperm DNA).

Visualization: Images were captured using a Nikon Microphot-SA microscope fitted with a cooled CCD camera (Digital Pixels, Brighton, UK) and analyzed with IPLab software (Scanalytics, Fairfax, VA).

RESULTS

In the transgenic line BLG/7, variable expression of the ovine β -lactoglobulin (the transgene) protein in milk reflects variegated transgene expression in mammary epithelial cells (DOBIE *et al.* 1996). This locus did not display age-related progressive silencing as the expression level for an individual mouse is stable between lactations (DOBIE *et al.* 1996). To determine the effect of backcrossing, the transgenic locus was bred on two different backgrounds for 13 generations. These mice were congenic, with ~99.4% of chromosomal loci being homozygous (excluding those loci closely linked to the transgene integration locus). Milk samples were collected from 18 BLG/7 CBA mice and 21 BLG/7 C57BL/6 mice. Nineteen milk samples were taken from a parental population of BLG/7 transgenic mice with a mixed CBA \times C57BL/6 genetic background, also bred for 13 generations.

Variegation occurs irrespective of genetic background:

We determined β -lactoglobulin protein levels on polyacrylamide gels by comparison to known β -lactoglobulin standards for the backcross and parental populations. Values were corrected for loading differences by comparison to mouse β -casein levels, which we have previously shown to be expressed uniformly in BLG/7 mice (DOBIE *et al.* 1996). Similar results were also apparent when serum albumin was used as the loading control protein (data not shown). Clearly, the BLG/7 transgene displays variable expression in both backcross populations (Figure 1). *In situ* hybridization for β -lactoglobulin expression in mammary tissue of backcross mice shows

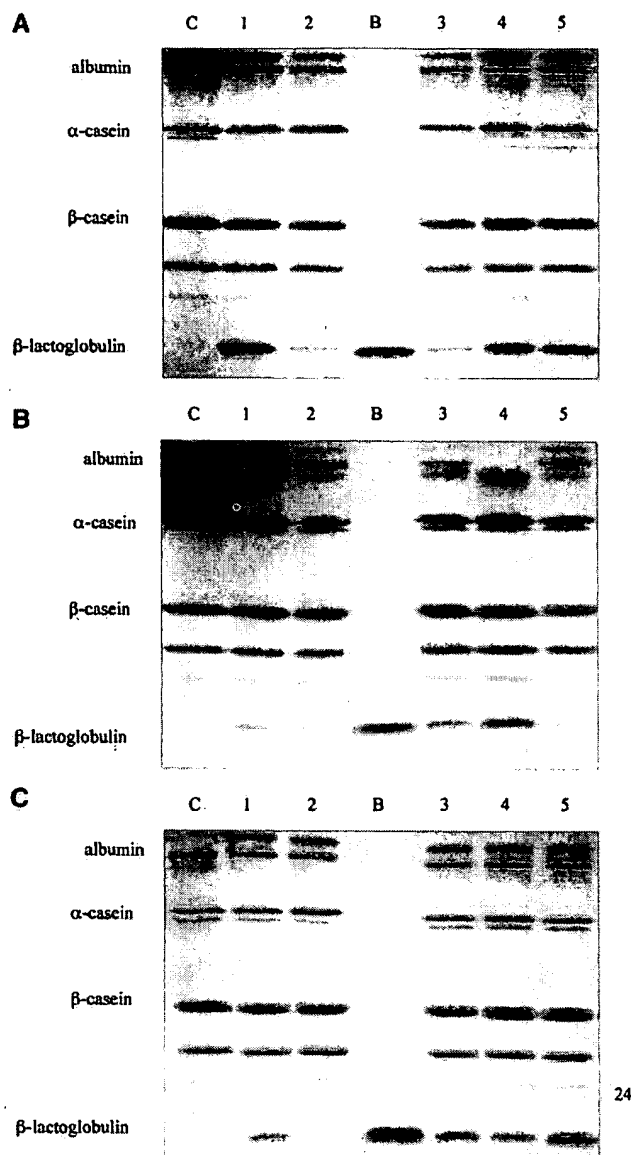


FIGURE 1.—BLG/7 locus displays variable expression. SDS/PAGE protein gels showing milk samples taken from 13th generation backcross CBA (A), C57BL/6 (B), and parental mice (C). Lane C, nontransgenic control milk; lane B, standard β -lactoglobulin sample: 2 μ g for A and B, 3 μ g for C. All samples were a 1/250 dilution of milk samples taken at day 11 of lactation from individual mice with 50 μ l (A and B) and 40 μ l (C) loaded per sample. Identities of milk proteins are indicated.

that the variable expression is due to variegation (Figure 2). The extent of variegation reflects β -lactoglobulin levels in milk for the parental population (M. L. OPSAHL, A. SPRINGBETT, R. LATHE, A. COLMAN, M. MCCLENAGHAN and C. WHITELAW, unpublished results) and for those backcross mice analyzed (data not shown). Expression level for individual mice from each backcross population and the parental population were determined (Figure 3) and the mean, standard deviation, and coefficient of variation for each population were calculated (Table 1).

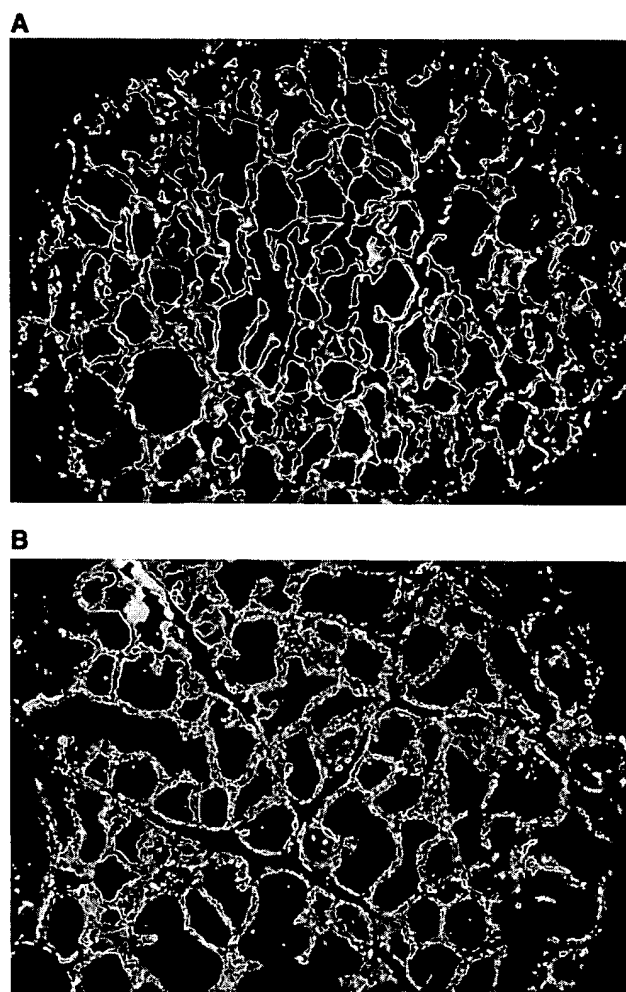


FIGURE 2.—BLG/7 locus displays variegated expression pattern. (A) FISH of backcross wax-embedded mammary tissue probed simultaneously for BLG (Cy3) and β -casein (fluorescein) expression. (B) Nontransgenic control tissue probed for BLG (Cy3) and β -casein (fluorescein). Note that residual wax fluoresces bright yellowish-white.

Expression of transgenic protein levels was compared. Several effects were observed. The variance of β -lactoglobulin protein levels in the C57BL/6 population was significantly greater than for the CBA population ($P < 0.05$). The mean for the C57BL/6 population was also significantly higher than for the CBA population ($P < 0.01$). Thus, although BLG/7 still variegates after extensive backcrossing into CBA or C57BL/6 genetic backgrounds, there are significant differences in expression levels between these populations. For both backcross populations, the mean and variance differed significantly from the parental population (both $P < 0.01$), with values considerably lowered in both cases. A mild trend toward a reduction of mean expression and variegation in the CBA third generation backcross (DOBIE *et al.* 1996) is significantly enhanced in the 13th generation population, and similar trends that were not apparent in the 3rd generation C57BL/6 background have now become evident. Therefore, modifiers that down-

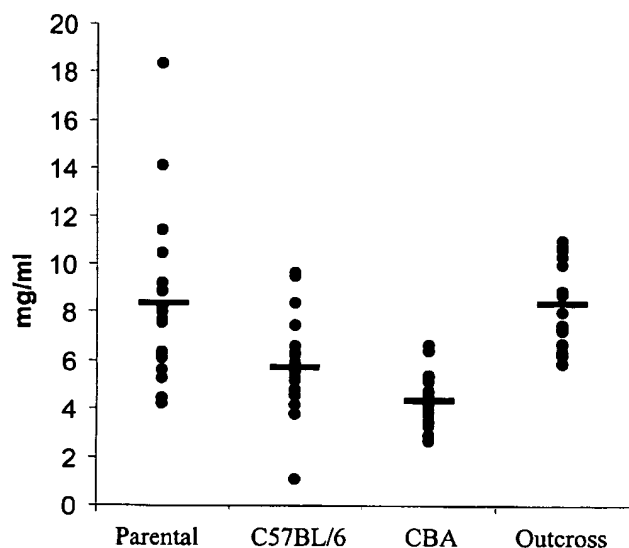


FIGURE 3.—Comparison of BLG/7 expression on different genetic backgrounds. Yield (milligrams per milliliter) of β -lactoglobulin from individual mice in the mixed parental, backcross C57BL/6, backcross CBA, and outcross populations. Each dot represents an average of three estimates for each individual mouse. Mean levels for each population are shown as a horizontal bar.

regulate without silencing expression of the variegating transgene exist. In addition, the maintenance of transgene expression levels in the parental population indicates that there is no progressive silencing of the transgene through repeated rounds of germline transmission.

The BLG/7 transgene is selectively sensitive to inbreeding depression effects: Inbreeding depression led to a slight reduction in total milk protein levels for both backcross populations (Table 2), with total protein content down by 15%. Unexpectedly, transgene expression was found to be depressed to a greater extent (down 29% in the C57BL/6 population and 47% in the CBA population). Thus, factors contributing to inbreeding depression may selectively suppress the transgene.

Outcrossing rescues β -lactoglobulin expression to parental levels: To determine if the inbreeding suppression of transgene expression could be maintained in a mixed genetic background we outcrossed one of the backcross populations (C57BL/6 BLG7 males to CBA,

generating first-generation outcross females). Variegation was still very much evident in these mice. The mean β -lactoglobulin expression level reverted to that of the parental BLG/7 population, *i.e.*, 8.2 mg/ml. This differed significantly from the backcross C57BL/6 BLG/7 population (Table 3; Figure 3). This further supports our conclusion that there is no progressive silencing of the transgene through repeated rounds of germline transmission when in a mixed genetic background.

The coefficient of variation was reduced for the outcross *vs.* parental mice (Table 3), although expression was always within the range seen for the parental population (Figure 3). Differences in genetic composition could account for this. The outcross mice are very nearly 50% CBA and 50% C57BL/6, while the parental mice are maintained on a mixed CBA \times C57BL/6 background through $F_2 \times F_1$ matings. Thus, individuals in the parental population can be homozygous (for CBA or C57BL/6) for a variable number of loci.

In conclusion, the reduced transgene expression seen in the backcross populations is not due to heritable change at the transgenic locus. Rather, this indicates that the inbreeding and background effects reflect recessive allele(s) with the manifestation of the allele(s) not being inherited.

DISCUSSION

We have investigated the influence of genetic background on expression of a known variegating transgene. Variable transgene expression levels were seen on both backgrounds studied, even after 13 generations of backcrossing. The effects of inbreeding depression were more severe for our transgene than endogenous milk genes. Outcrossing resulted in complete reversion to parental expression levels.

Modifiers of variegation: The BLG/7 locus displays variable expression in both 13th generation CBA and C57BL/6 backcross populations. Therefore, the variegation observed is an inherent property of the BLG/7 transgene array at its site of integration—a position effect. This epigenetic property is inherited through generations. Variegating gene expression has been reported in many different organisms (CATTANACH 1974; HENIKOFF 1990; MARTIN and WHITELAW 1996; DOBIE *et al.* 1997). Apart from those that are the result of mutations, most examples derive from translocation or transgene insertion into unfavorable positions in the genome. These positions are often close to heterochromatic chromatin, and the cause of the subsequent mosaic expression pattern is believed to be a variable spread of heterochromatic proteins into the inserted gene array (PARO and HOGNESS 1991). Modifiers of PEV have been identified, many being components of chromatin (SINGH 1994), and there is evidence that the mammalian homologs can modify expression of a variegating gene (STAN-

TABLE 1

Analysis of expression in 13th generation backcross and parental populations

| Genotype | Mean (mg/ml) | n | SD | Coefficient of variation (%) |
|----------------|--------------|----|------|------------------------------|
| CBA BLG/7 | 4.3 | 18 | 1.17 | 27.3 |
| C57BL/6 BLG/7 | 5.8 | 21 | 1.92 | 33.1 |
| Parental BLG/7 | 8.2 | 19 | 3.46 | 42 |

TABLE 2
Micro-Kjeldahl analysis of total milk protein content

| Genotype | β -Lactoglobulin ^a | Total protein ^a | Protein minus β -lactoglobulin ^a | % of endogenous proteins ^b | % reduction in β -lactoglobulin ^b |
|----------------|-------------------------------------|----------------------------|---|---------------------------------------|--|
| CBA BLG/7 | 4 | 85 | 81 | 85 | 47 |
| C57BL/6 BLG/7 | 6 | 87 | 81 | 85 | 29 |
| Parental BLG/7 | 8 | 103 | 95 | 100 | 0 |

^a Protein in milligrams per milliliter.

^b Relative to parental BLG/7 values.

KUNAS *et al.* 1998; AAGAARD *et al.* 1999; FESTENSTEIN *et al.* 1999; McMORROW *et al.* 2000). Presumably some aspect of the BLG/7 transgene locus renders it a target for this type of nuclear factor.

Evidence for strain-specific major modifier locus effects in mice has been found (ALLEN *et al.* 1990; SCHWEIZER *et al.* 1998; MAYEUX-PORTAS *et al.* 2000). These reported strain differences are generally evident by the second or third generation backcross, with clear segregation patterns emerging that are indicative of a single locus. In addition, overexpression of a single protein can induce a major modification of expression in a variegating transgene (FESTENSTEIN *et al.* 1999; McMORROW *et al.* 2000). For the BLG/7 locus, we concluded previously from the third generation backcross that no single major modifier is present (DOBIE *et al.* 1996). Nevertheless, in the 13th generation backcrosses we see both a lowering in the mean β -lactoglobulin level and an accompanying reduction in the variability of transgene expression. This suggests that rather than a single major modifier affecting the BLG/7 transgene, multiple loci exist, each having a subtle but cumulative effect on expression. There are also differences between the two backcross populations, implying that the modifiers differ in either composition or concentration in the CBA and C57BL/6 backgrounds.

Inbreeding depression selectively downregulates transgene expression: As a consequence of inbreeding depression the total milk protein content of both backcrosses was slightly reduced compared to the parental milks. We additionally observed a marked selective reduction of β -lactoglobulin levels with respect to that of endogenous milk proteins. We believe this is the first reported case of selective "discrimination" of a trans-

gene in an inbred environment. Upregulation of β -lactoglobulin upon outcrossing rules out the presence of sequence mutations or heritable epigenetic modifications at the β -lactoglobulin locus.

We propose that, in contrast to the endogenous mouse milk proteins, the BLG/7 transgene locus is specifically susceptible to inbreeding depression. It is not clear what aspect of the BLG/7 locus confers this increased sensitivity. The transgene may lack regulatory regions present at the endogenous milk protein gene loci that provide a compensatory mechanism. Alternatively, the multicopy nature of the transgene array may be recognized in some way (DORER and HENIKOFF 1994; MATZKE and MATZKE 1995; HSIEH and FIRE 2000). Whatever the cause, it presumably reflects a recruiting disadvantage for the β -lactoglobulin transgene sequences with regard to the endogenous genes. In support of this being a general property of transgenes, the overexpression of the suppressive heterochromatic protein M31 (the mammalian homolog of the *Drosophila* heterochromatic protein 1) in transgenic mice selectively silenced a transgene without causing any phenotypic changes (FESTENSTEIN *et al.* 1999). Subsequently, and again without reported phenotypic changes, dosage-dependent effects of enhancers and suppressors of variegation were observed in a transgene (McMORROW *et al.* 2000). As heterochromatic factors are involved in the organization of higher-order chromatin, it is tempting to speculate that the BLG/7 transgene locus is compromised with regard to its spatial organization within the nucleus. Furthermore, because variable expression is inherited while inbreeding suppression is not, some differences in the mechanisms leading to these two forms of silencing must exist.

It has been proposed that the extreme consequence of inbreeding depression could be that a gene is transcriptionally silenced (MATZKE *et al.* 1993). We did not observe such total silencing of expression for the BLG/7 transgene, only downregulation of expression in both genetic backgrounds. Should this phenomenon prove to be common, then current backcrossing strategies designed to alleviate unstable transgene expression may exacerbate the problem instead of curing it.

Genome policing: In plants it has been proposed that

TABLE 3

Analysis of expression in parental and outcross populations

| Genotype | Mean (mg/ml) | n | SD | Coefficient of variation (%) |
|----------------|--------------|----|------|------------------------------|
| Parental BLG/7 | 8.2 | 19 | 3.46 | 42 |
| Outcross BLG/7 | 8.2 | 16 | 1.80 | 22 |

a genome defense system operates to inactivate foreign or invasive sequences such as transgenes and transposable elements (MATZKE and MATZKE 1998). One aspect of this policing of the plant genome is that the extent of silencing increases through generations (ASSAAD *et al.* 1993; GUO *et al.* 1998) with similar effects reported in insects (JENSEN *et al.* 1999) and fish (AMSTERDAM *et al.* 1995). We are unaware of any reported generation-associated silencing of transgenes in mammals. In our study, in contrast to the selective silencing of the β -lactoglobulin transgene during backcrossing, we saw no evidence for a generation-associated progressive silencing of our variegating transgene when present in a mixed genetic background. Indeed, expression of this transgene has remained essentially constant since its generation in 1987 (SIMONS *et al.* 1987). Therefore, if a genome-policing mechanism exists in mammals to selectively silence foreign sequences, then some transgenes, *e.g.*, BLG/7, can evade detection.

We are grateful to the Small Animals Unit staff for care and maintenance of the mice, which was carried out under HO (UK) regulations. We also thank Caroline McCorquedale and George Russell for constructive advice on use of the Multi Analyst and data processing. Many thanks also to Eleanor Noble (Hannah Research Institute) for kindly agreeing to do the micro-Kjeldahl analysis of our milk samples. This project was initiated by Maggie McClenaghan. Sarah Reid was an Honours Biochemistry undergraduate at the University of Edinburgh. The work was supported by the Biotechnology and Biological Sciences Research Council through a CASE award with PPL Therapeutics to Margaret Opsahl.

LITERATURE CITED

- AAGAARD, L., G. LAIBLE, P. SELENKO, M. SCHMID, R. DORN *et al.*, 1999 Functional mammalian homologues of the *Drosophila* PEV-modifier *Su(var)3-9* encode centromere-associated proteins which complex with the heterochromatin component M31. *EMBO J.* 18: 1923–1938.
- ALLEN, N. D., M. L. NORRIS and M. A. SURANI, 1990 Epigenetic control of transgene expression and imprinting by genotype-specific modifiers. *Cell* 61: 853–861.
- AMSTERDAM, A., S. LIN and N. HOPKINS, 1995 The *Aequorea victoria* green fluorescent protein can be used as a reporter in live zebrafish. *Dev. Biol.* 171: 123–129.
- ASCHAFFENBURG, R., and J. DREWRY, 1957 Improved method for the preparation of crystalline β -lactoglobulin and α -lactalbumin from cow's milk. *Biochem. J.* 65: 273–277.
- ASSAAD, F. F., K. L. TUCKER and E. R. SIGNER, 1993 Epigenetic repeat-induced gene silencing (RIGS) in *Arabidopsis*. *Plant Mol. Biol.* 22: 1067–1085.
- CATTENACH, B. M., 1974 Position effect variegation in the mouse. *Genet. Res.* 23: 291–306.
- DOBIE, K. W., M. LEE, J. A. FANTES, E. GRAHAM, A. J. CLARK *et al.*, 1996 Variegated transgene expression in mouse mammary gland is determined by the transgene integration locus. *Proc. Natl. Acad. Sci. USA* 93: 6659–6664.
- DOBIE, K. W., M. METHALI, M. MCCLENAGHAN and R. LATHE, 1997 Variegated gene expression in mice. *Trends Genet.* 13(4): 127–130.
- DORER, D. R., and S. HENIKOFF, 1994 Expansions of transgene repeats cause heterochromatin formation and gene silencing in *Drosophila*. *Cell* 77: 993–1002.
- ELLIOT, J. I., R. FESTENSTEIN, M. TOLAINI and D. KIOUSSIS, 1995 Random inactivation of a transgene under the control of a hybrid hCD2 locus control region/Ig enhancer regulatory element. *EMBO J.* 14: 575–584.
- FESTENSTEIN, R., M. TOLAINI, P. CORBELLA, C. MAMALAKI, J. PARRINGTON *et al.*, 1996 Locus control region function and heterochromatin-induced position effect variegation. *Science* 271: 1123–1125.
- FESTENSTEIN, R., S. SHARGHI-NAMINI, M. FOX, K. RODERICK, M. TOLAINI *et al.*, 1999 Heterochromatin protein 1 modifies mammalian PEV in a dose- and chromosomal-context-dependent manner. *Nat. Genet.* 23: 457–461.
- GIRALDO, P., E. GIMENEZ and L. MONTOLIU, 1999 The use of yeast artificial chromosomes in transgenic animals: expression studies of the tyrosinase gene in transgenic mice. *Genet. Anal.* 15: 175–178.
- GUO, H. S., M. T. CERVERA and J. A. GARCIA, 1998 Plum pox potyvirus resistance associated to transgene silencing that can be stabilized after different number of plant generations. *Gene* 206: 263–272.
- HENIKOFF, S., 1990 Position-effect variegation after 60 years. *Trends Genet.* 6: 422–426.
- HSIEH, J., and A. FIRE, 2000 Recognition and silencing of repeated DNA. *Annu. Rev. Genet.* 34: 187–204.
- JENSEN, S., M. P. GASSAMA and T. HEIDMANN, 1999 Cosuppression of I transposon activity in *Drosophila* by I-containing sense and antisense transgenes. *Genetics* 153: 1767–1774.
- MARTIN, D. I., and E. WHITELAW, 1996 Vagaries of variegating transgenes. *Bioessays* 18: 919–923.
- MATZKE, M. A., and A. J. M. MATZKE, 1995 Homology-dependent gene silencing in transgenic plants: What does it really tell us? *Trends Genet.* 11: 1–3.
- MATZKE, M. A., and A. J. M. MATZKE, 1998 Gene silencing in plants: relevance for genome evolution and the acquisition of genome methylation patterns. *Epigenetics* 214: 168–180.
- MATZKE, M. A., F. NEUHUBER and A. J. MATZKE, 1993 A variety of epistatic interactions can occur between partially homologous transgene loci brought together by sexual crossing. *Mol. Gen. Genet.* 236: 379–386.
- MAYEUX-PORTAS, V., S. E. FILE, C. L. STEWART and R. J. MORRIS, 2000 Mice lacking the cell adhesion molecule Thy-1 fail to use socially transmitted cues to direct their choice of food. *Curr. Biol.* 10: 68–75.
- MCCLENAGHAN, M., A. SPRINGBETT, R. M. WALLACE, C. J. WILDE and A. J. CLARK, 1995 Secretory proteins compete for production in the mammary gland of transgenic mice. *Biochem. J.* 310: 637–641.
- MCCOWAN, R., B. VAMPBELL, A. PETERSON and C. SAPIENZA, 1989 Cellular mosaicism in the methylation and expression of hemizygous loci in the mouse. *Genes Dev.* 268: 1669–1676.
- McMORROW, T., A. VAN DEN WIJNGAARD, A. WOLLENSCHLAGER, M. VAN DE CORPUT, K. MONKHORST *et al.*, 2000 Activation of the β -globin locus by transcription factors and chromatin modifiers. *EMBO J.* 19: 4986–4996.
- MULLER, H. J., 1930 Types of visible variations induced by X-rays in *Drosophila*. *J. Genet.* 22: 299–334.
- PARO, R., and D. S. HOGNESS, 1991 The Polycomb protein shares a homologous domain with a heterochromatin-associated protein of *Drosophila*. *Proc. Natl. Acad. Sci. USA* 88: 263–267.
- ROBERTSON, G., D. GARRICK, M. WILSON, D. I. K. MARTIN and E. WHITELAW, 1996 Age-dependent silencing of globin transgenes in the mouse. *Nucleic Acids Res.* 24: 1465–1471.
- SCHWEIZER, J., P. VALENZE-SCHAERLY, F. GORET and C. POURCEL, 1998 Control of expression and methylation of a hepatitis B virus transgene by strain-specific modifiers. *DNA Cell Biol.* 17: 427–435.
- SIMONS, J. P., M. MCCLENAGHAN and A. J. CLARK, 1987 Alteration of the quality of milk by expression of sheep β -lactoglobulin in transgenic mice. *Nature* 328: 530–532.
- SINGH, P. B., 1994 Molecular mechanisms of cellular determination: their relation to chromatin structure and parental imprinting. *J. Cell Sci.* 107: 2653–2668.
- STANKUNAS, K., J. BERGER, C. RUSE, D. A. R. SINCLAIR, F. RANDAZZO *et al.*, 1998 The *Enhancer of Polycomb* gene of *Drosophila* encodes a chromatin protein conserved in yeast and mammals. *Development* 125: 4055–4066.
- SUTHERLAND, H. G., M. KEARNS, H. D. MORGAN, A. P. HEADLEY, C. MORRIS *et al.*, 2000 Reactivation of heritably silenced gene expression in mice. *Mamm. Genome* 11: 347–355.
- WHITELAW, C. B. A., S. HARRIS, M. MCCLENAGHAN, J. P. SIMONS and A. J. CLARK, 1992 Position independent expression of the ovine β -lactoglobulin gene in mice. *Biochem. J.* 286: 31–39.

Communicating editor: S. HENIKOFF